# Restoring ancestral phenotypes is a general pattern in gene expression evolution during adaptation to new environments in Tribolium castaneum

Koch, Eva L ; Guillaume, Frederic

Abstract: Plasticity and evolution are two processes allowing populations to respond to environmental changes, but how both are related and impact each other remains controversial. We studied plastic and evolutionary responses in gene expression of Tribolium castaneum after exposure of the beetles to new environments that differed from ancestral conditions in temperature, humidity or both. Using experimental evolution with 10 replicated lines per condition, we were able to demonstrate adaptation after 20 generations. We measured whole-transcriptome gene expression with RNA-sequencing to infer evolutionary and plastic changes. We found more evidence for changes in mean expression (shift in the intercept of reaction norms) in adapted lines than for changes in plasticity (shifts in slopes). Plasticity was mainly preserved in selected lines and was responsible for a large part of the phenotypic divergence in expression between ancestral and new conditions. However, we found that genes with the largest evolutionary changes in expression also evolved reduced plasticity and often showed expression levels closer to the ancestral stage. Results obtained in the three different conditions were similar, suggesting that restoration of ancestral expression levels during adaptation is a general evolutionary pattern. With a larger sample in the most stressful condition, we were able to detect a positive correlation between the proportion of genes with reversion of the ancestral plastic response and mean fitness per selection line.

DOI: https://doi.org/10.1111/mec.15607

MS. EVA LOUISA KOCH (Orcid ID : 0000-0001-8366-4897)

DR. FRÉDÉRIC GUILLAUME (Orcid ID : 0000-0003-0874-0081)

Article type : Original Article

# Restoring ancestral phenotypes is a general pattern in gene expression evolution during adaptation to new environments in *Tribolium castaneum*

Eva L. Koch[1,2] and Frédéric Guillaume[1]

[1] Department of Evolutionary Biology and Environmental Studies, University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland

[2] Department of Animal and Plant Science, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom



This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/MEC.15607

Corresponding author:
Frédéric Guillaume
phone: +41 44 635 66 23
frederic.guillaume@ieu.uzh.ch

**Abstract**

Plasticity and evolution are two processes allowing populations to respond to environmental changes, but how both are related and impact each other is still controversial. We studied plastic and evolutionary responses in gene expression of *Tribolium castaneum* after beetles' exposure to new environments that differed from ancestral conditions in temperature, humidity or both. Using experimental evolution with ten replicated lines per condition, we were able to demonstrate adaptation after 20 generations. We measured whole-transcriptome gene expression with RNA-seq to infer evolutionary and plastic changes. We found more evidence for changes in mean expression (shift in the intercept of reaction norms) in adapted lines than for changes in plasticity (shifts in slopes). Plasticity was mainly preserved in selected lines and was responsible for a large part of the phenotypic divergence in expression between ancestral and new conditions. However, we found that genes with the largest evolutionary changes in expression also evolved reduced plasticity and often showed expression levels closer to the ancestral stage. Results obtained in the three different conditions were similar, suggesting that restoration of ancestral expression levels during adaptation is a general evolutionary pattern. With a larger sample in the most stressful condition, we were able to detect a positive correlation between proportion of genes with reversion of the ancestral plastic response and mean fitness per selection line.

**Introduction**

Whenever facing environmental change, populations can adapt to new phenotypic optima by plasticity and evolution. Plasticity is the ability of a single genotype to produce multiple phenotypes as a function of the environment. It is often seen as an immediate response of individuals to changes in their environment. In contrast, evolution requires a change in allele frequencies within a population. This process occurs over several generations and represents a more long-term response, which can result in local adaptation. It is still not well understood how these two processes are related and interact with each other (de Jong, 2005; Forsman, 2015; Ghalambor, McKay, Carroll, & Reznick, 2007; Price, Qvarnström, & Irwin, 2003; Wund, 2012).

By changing the distribution of phenotypes on which selection can act, plasticity interferes with the process of evolution in a population (de Jong, 2005; Pfennig et al., 2010; Massimo Pigliucci, 2005; Price et al., 2003). If plasticity allows the population to perfectly match its new phenotypic optimum, it will prevent selection and thus evolution (Ghalambor et al., 2007). On the other hand, plasticity is also crucial for a population's persistence and can reduce the costs of selection (Chevin, Lande, & Mace, 2010; Pavey,

Collin, Nosil, & Rogers, 2010). It can prevent extinction and protect populations from bottleneck effects, thereby maintaining a higher genetic variation on which subsequently selection can act (Fitzpatrick, 2012; Pfennig et al., 2010; Massimo Pigliucci, 2005). There is both theoretical (Chevin et al., 2010; Draghi & Whitlock, 2012; Fierst, 2011) and empirical work (Schaum, Rost, Millar, & Collins, 2013) demonstrating that more plastic populations exhibit faster evolution. The benefits of plasticity for persisting in new habitats were also demonstrated in invasive species (Molina-Montenegro, Peñuelas, Munné-Bosch, & Sardans, 2012; Pichancourt & van Klinken, 2012; Yeh & Price, 2004).

Yet, plasticity is also a trait that can evolve during population adaptation to new or varying environments (Draghi & Whitlock, 2012; Gavrilets & Scheiner, 1993; Lande, 2009; Schmid, Dallo, & Guillaume, 2019; Via & Lande, 1985). The extent of plasticity can be represented as a reaction norm (Scheiner, 1993), which is the phenotypic trait value as a function of an environmental variable. Evolution can affect the reaction norm in two ways: The intercept can be shifted, corresponding to a change in the mean phenotypic value, or the slope of the reaction norm, i.e. the plasticity, can be changed. Thus, natural selection may act on the two underlying traits defining reaction norms, the mean trait value and its plasticity, provided there is sufficient genetic variation in reaction norms (Garland & Kelly, 2008; Nussey, Postma, Gienapp, & Visser, 2005).

The evolutionary dynamics of reaction norms depend on the relationship between plastic responses and local adaptation in a population. Plastic responses are adaptive when they increase the fitness of an individual. If plastic responses are adaptive, but not sufficient to reach the phenotypic optimum of the plastic trait, evolution should work in the same direction as the plastic response (referred to as Baldwin effect (Crispo, 2007) or cogradient variation (Conover, Duffy, & Hice, 2009). In this case, selection may favour the most plastic individuals, causing evolved populations to exhibit a higher plasticity than their ancestors (Crispo 2007; Lande 2009). Another possible outcome is genetic assimilation: An initially environmentally induced phenotypic change can become fixed in the population by a loss of plasticity and be continuously expressed even in the ancestral environment (Levis & Pfennig, 2016; Pigliucci, Murren, & Schlichting, 2006). In contrast, if plastic responses are maladaptive, i.e. decrease individual fitness, we expect to observe evolutionary changes opposite to plasticity (countergradient variation (Conover et al., 2009) or genetic compensation (Grether 2005)). Maladaptive plasticity was proposed as a possible mechanism promoting evolution since it moves phenotypes further away from their optimum and thereby increases the strength of selection on the phenotypes (Ghalambor et al., 2007). Both co-gradient (Barton, Sunnucks, Norgate, Murray, & Kearney, 2014; Conover et al., 2009) and counter-gradient

(Conover et al., 2009; Ghalambor et al., 2015; Laugen, Laurila, Räsänen, & Merilä, 2003) evolutionary changes have been found, indicating that adaptive and maladaptive plasticity are common. Reversion of ancestral plasticity occurs more frequently (Ho & Zhang, 2018), indicating that plastic responses are often not beneficial for long-term adaptation.

Plastic responses in physiology, behaviour or morphological traits are often initiated by changes in gene expression (Hodgins-Davis & Townsend, 2009; Wray, 2007). The transcriptome represents a direct link between genotype and phenotype making it particularly interesting to study the interplay between plasticity and evolution. Transcription is highly plastic and modulating expression levels is an important part of an organism's physiological adjustment to environmental change (Gibson, 2008; McCairns & Bernatchez, 2009). On the other hand, there are also many studies demonstrating evolutionary divergence in gene expression between locally adapted populations (Alvarez, Schrey, & Richards, 2015; Guo et al., 2016; Romero, Ruvinsky, & Gilad, 2012; Townsend, Cavalieri, & Hartl, 2003; Whitehead & Crawford, 2006). Gene expression may even evolve more rapidly than changes in proteins since mutations affecting the magnitude of expression are less likely to be deleterious than changes in protein structures (Carroll, 2005; Wray, 2007). However, it is not clear how fast plasticity in gene expression can change. Some studies reported changes in plasticity in few genes after adaptation to new conditions (Morris et al., 2014; Passow et al., 2017; von Heckel, Stephan, & Hutter, 2016), whereas others found only limited evolution of plasticity (Yampolsky, Glazko, & Fry, 2012) or less than expected (Huang & Agrawal, 2016).

In our study, we used whole transcriptomes to understand the interplay between plasticity and evolution at the gene expression level during adaptation to new environments. We used the model organism *Tribolium castaneum* (red flour beetle) in an experimental evolution approach. We submitted the beetle to three stressful environments for 20 generations using a high number of evolutionary replicates in a controlled laboratory setting. We used ten replicate selection lines per condition and measured lines' adaptation to treatment conditions in a large fitness assay. We sequenced a total of more than 200 individuals across all conditions. This gave us high statistical power to evaluate the relative contribution of plastic and evolutionary changes to the total divergence in gene expression among conditions. More specifically, we were interested to test whether the same genes exhibited both evolutionary and plastic changes in a new environment and whether evolved changes were in the same direction as their ancestral plasticity. In doing so, we could test how plasticity affected evolution, and whether plasticity itself evolved.

## Material and Methods

*Animal rearing, experimental evolution*

We used the *Tribolium castaneum* Cro1 strain (Milutinović, Stolpe, Peuß, Armitage, & Kurtz, 2013), collected from a wild population in 2010 and adapted to lab standard conditions (33°C, 70% relative humidity (r.h.)) for more than 20 generations. Beetles were kept in 24h darkness on organic wheat flour mixed with 10% organic baker's yeast. We sterilized flour and yeast by heating them for 12h at 80°C before use. To test for adaptation to new environmental conditions we used replicate lines and exposed them to three treatment and Control (CT) conditions. The conditions in the treatments were: Dry (D): 33°C and 30% r. h.; Hot (H): 37°C and 70% r. h.; Hot-Dry (HD): 37°C and 30% r. h. To generate replicate lines, we used 120 individuals (60 females and 60 males in the pupal stage) and placed them into a vial containing 80g medium. We produced six lines per selection regime (treatments plus control), resulting in a total of 24 lines. For each new generation, we randomly collected 120 pupae and placed them into a new vial. After seven to ten days, in which the pupae became adults, mated and laid eggs, adult beetles were removed by sieving the medium. We waited until the next generation (eggs/larvae in the medium) had reached the pupal stage and again collected 120 pupae per line to establish the next generation. This is similar to natural selection since individuals, depending on their fitness, do not contribute equally to the next generation. In generation 15 we produced additional mixed lines to prevent loss of genetic diversity by gene drift and inbreeding, which might impede adaptation: We mixed the six replicate lines of each selection regime in equal proportions (20 individuals from each replicate line) four times, resulting in four mixed lines with 120 individuals each. In total we had 39 lines: six normal and four mixed lines per selection regime (one line in D became extinct). The transplant experiment to test for adaptation was conducted in generation 22.

*Reciprocal transplant and fitness assay*

Before testing for adaptation, all lines stayed for two generations in the same condition to reduce potential maternal or epigenetic effects (Supporting information, Figure S1): Beetles of generation 20 from all selection lines were transferred to control conditions, in which they stayed for one week to mate and lay eggs. After removal of the adults, we waited until their offspring had reached the pupal stage and separated males and females. These individuals (generation 21) developed completely in control conditions. When they reached the adult stage, we created 13 full-sib families per selection line by transferring one virgin male with one virgin female of the same selection line in 15mL tubes with 1g of

medium. After four days, in which the beetles could mate and lay eggs, 9g of medium was added to provide food for the developing offspring and each mating pair was transferred to a new vial. We repeated this three times, resulting in four vials per mating pair containing medium and eggs. Immediately after removal of the mating pair, vials of each mating pair were randomly assigned to the four different conditions, resulting in full-sib families split across all conditions. These beetles were transferred to the treatments at the egg stage. As soon as offspring in these vials had reached the pupal stage, males and females (four females and four males per family and condition) were separated and transferred to 15 mL tubes with 5 g of medium and remained there until they were used for the fitness assay two weeks later. They developed completely in treatment conditions. We then assessed their performance in each condition by estimating their fitness to test for adaptation. A virgin male and a virgin female of the same selection line from the same condition but from different families were again placed into a 15 mL tube with 1 g medium. After four days, the mating pair was removed. Males and females were transferred to 1 mL Eppendorf tubes (one individual per tube), immediately frozen in liquid nitrogen and stored at -80°C to use them for gene expression measurements. 9 g medium was added to the mating tube. After four weeks (in Control and Hot) or five weeks (Dry and Hot-Dry), all offspring had reached the adult stage and were counted. We used the number of adult offspring as an estimate of the fitness of a mating pair. Sample size for each line and treatment can be found in Supporting information Table S1.

*Statistical analysis*

To test whether selection regime significantly influenced number of offspring produced and test whether 20 generations in the treatments resulted in adaptation, we compared offspring numbers of selection lines in their native condition to Control lines transferred to the same condition. We applied linear mixed models using the R-packages *lme4* (Bates, Mächler, Bolker, & Walker, 2015), and *lmertest* (Kuznetsova, Brockhoff, & Christensen, 2017) and *lsmeans* (Lenth, 2016) to obtain *p*-values and confidence intervals. We included line and family as random factors, selection and line type (mixed/normal) and their interaction as fixed effects. To test whether the selection regime influenced how lines responded to the treatments, we used a linear mixed model with offspring number in control and treatment conditions as response variable, condition, selection regime, line type (normal/mixed) and interactions as fixed effects and line, family and interaction between line and condition as random effects. A significant interaction between condition and selection regime indicates a significant effect of the selection regime (evolution) on the response to the conditions (plastic response).

*RNA extraction, library preparation and sequencing*

208 female beetles (Table 1, for sample sizes per line and treatment see Table S2) stored at -80°C were homogenized in Tri-Reagent® (Zymo Research, California, USA) using an electric bead mill. RNA was extracted with the RNA Mini Prep kit (Zimo Research, California, USA) following the instructions of the manufacturer. RNA-quality was checked on a TapeStation (Agilent, Waldbronn, Germany) and concentrations were measured with aQubit® Fluorometer (Life Technologies, California, USA). Libraries were created with 500 ng RNA for each individual separately with the LEXOGEN mRNA-Seq Library Kit following the manual (LEXOGEN GmbH, Vienne, Austria). Library quality was checked on a TapeStation (Agilent, Waldbronn, Germany) and concentrations were determined by qPCR. Libraries were diluted to the same molarity and pooled (33-36 libraries per pool). All treatments and selection regimes were randomized during RNA-extraction, library preparation, and sequencing. Single-end sequencing was performed in five runs on the Illumina NextSeq 500 (Illumina, Inc, California, USA) using the 75 cycles High Output Kit. After quality control using FastQC (www.bioinformatics.bbsrc.ac.uk/projects/fastqc) reads (adaptors were trimmed and the first 10 bases were hard trimmed, minimum average quality Q10, minimum tail quality 10, minimum read length 20) were mapped against the reference genome (ftp://ftp.ensemblgenomes.org/pub/release30/metazoa/gtf/tribolium_castaneum/Tribolium_castaneum.Tcas3.30.gtf.gz) with STAR v.2.5 (Dobin et al., 2013). We then used FeatureCounts (Liao, Smyth, & Shi, 2014) to count the number of reads that mapped to each gene in the reference genome. Mapping as well as read counting was performed within the data analysis framework SUSHI (Hatakeyama et al., 2016). We obtained expression data for 17078 genes.

*Gene expression analysis*

Gene expression analysis was done in R (R Core Team, 2017). We used the R package *edgeR* (Robinson, McCarthy, & Smyth, 2010) for normalizing (method: TMM) expression data to cpm (counts per million) after filtering lowly expressed genes (minimum of one cpm in at least two samples). For subsequent differential expression analysis we used the R package *limma* (Law, Chen, Shi, & Smyth, 2014; Ritchie et al., 2015). We treated lines from the same selection regime as evolutionary replicates (block effect in *limma*). We sequenced four individuals per selection line and treatment (see Table 1). We accounted for non-independence of individuals from the same line by using the *duplicateCorrelation* function in *limma*. This is equivalent to including line as random effect in a linear mixed model (Law et al., 2014). We then analysed each condition separately (e.g. Control-lines and Hot-Dry-lines in control and hot-dry condition) and tested for the effect of treatment and selection regime as well as their interaction while correcting for

batch effects (sequencing runs). A gene is classified as differentially expressed (DE) with a FDR ≤ 5% after adjusting for multiple testing (Benjamini and Hochberg 1995).

For control and hot-dry conditions where we also sequenced individuals from the mixed lines (see also Table S2) we could not detect differences in expression levels between mixed and normal lines. MDSplots (Supporting information Figure S3) did not show a clustering by line type. We therefore did not analyse them separately but included them in the analyses.

From the differential expression analysis, we obtained the number of differentially expressed genes (DE genes) within lines between conditions (plastic changes, see Figure S1) or between lines of different origins (Control vs. selection) within conditions (evolutionary changes, see Figure S1). The total phenotypic divergence in gene expression between Control and treatments (i.e. total change TC) is the differential expression (log2-fold change) between Control-lines in control condition and selection lines in the treatments (Figure 1A-C and Supporting information Figure S1). The ancestral plasticity ($PC_{CT}$) is the differential expression of Control-lines between control and treatment conditions, while the evolved plasticity, plasticity of the selection-lines, ($PC_{Sel}$) is the same difference measured in selection lines. The evolutionary changes are $EC_T$ when measured as differential expression between Control- and selection-lines in the treatments and $EC_{CT}$ when measured in Control (Figure 1). Finally, differences between plastic responses of control and selection lines (the interaction between condition and selection regime) give the evolutionary change in plasticity.

To test for significant effects of the selection regime on differential expression (i.e., on number of DE genes, log2-fold change, and correlations), we used a permutation test. We randomly assigned samples and their transcriptomes to either Control-selection or treatment (Dry, Hot, Hot-Dry) selection (number of samples for each selection was not changed) and repeated the DE analysis. We kept the original assignment to lines and conditions and repeated the DE analysis for each permuted data set. Observed values (e.g. number of DE genes, correlations) were considered significant if higher than the most extreme 5% of the distribution calculated from permutations.

To partition total divergence (difference between Control-lines in control and selection-lines in treatment) into changes explained by ancestral plasticity and evolutionary changes ($PC_{CT}$ and $EC_T$, see Figure 1A-C), we calculated the relative contribution of each component to the total. We used the log2-fold change of each gene to evaluate and compare the magnitude of the plastic and evolutionary changes (see also Stoks et al. 2016). In a second step, we used the normalized read counts (cpm, TMM-normalized) corrected for batch effects (sequencing runs) using the *removeBatchEffect* function in the *limma* R package (Ritchie et al., 2015). We used these counts to quantify expression levels in Control and treatments. This allowed us

to conduct an analysis for each selection lines separately as was described by Ho & Zhang (2018) (details see below).

*Comparing plasticity and evolution*

To infer the relationship between ancestral plasticity (plastic response of Control-lines) and evolution, we compared the direction of ancestral plastic responses to the direction of evolutionary changes. Evolution may *reinforce* the plastic response when the ancestral plastic response is in the same direction as the evolutionary change (Figure 1B). If the evolutionary change is in opposite direction, it *reverses* the ancestral plastic response (Figure 1A). To test which of these patterns was more prevalent, we followed Ho and Zhang (2018). Expression levels (cpm) of Control-lines in control conditions represented the original expression level (*Lo*), Control-lines in treatment the ancestral plastic expression level (*Lp*), and selection lines in their respective condition the adaptive expression level (*La*) (Figure 1). For subsequent analysis we used genes with appreciable ancestral plasticity ($|Lp - Lo| > 0.2\ Lo$) and evolutionary response ($|La - Lp| > 0.2\ Lo$) (Ho & Zhang, 2018) and calculated the proportion of plastic genes with reinforced or reversed changes. To confirm that our results were not sensitive to the applied cutoff (20% of original expression levels *Lo*), we repeated the analysis with a cutoff of 50% of *Lo* and without any cutoff (Supporting information, Figure S5 G-L). It was pointed out (Mallard et al. 2018; Ho and Zhang 2019) that an excess of reversions relative to reinforcements is expected to be observed due to a statistical artefact that cannot be completely removed by permutation tests. Both, evolutionary as well as plastic responses rely on expression levels of Control-lines in treatment conditions (*Lp* in Figure 1). To obtain independent measures for *Lp*, we split the Control-lines randomly in two groups. We used one of them as reference in treatment conditions to infer evolutionary changes (differences between selection-lines and Control-lines), and the second group for measuring the ancestral plastic response (differences in expression levels of Control-lines between control and treatment conditions). Thus, we avoided any confounding effects of using the same measure (Control-levels in treatment) to infer plastic and evolved responses. As an alternative way to avoid the problem of non-independence between plastic and evolved responses, Ho and Zhang (2019) proposed a parametric bootstrap. This approach has the advantage that it does not suffer from the reduction in statistical power that inevitably results from splitting Control-lines into two groups. Results are shown in Supporting information, Figure S5 D-F. In addition, we compared our results to a conservative null distribution of proportion of reversion when randomly categorizing half of the control-lines as evolved (see Supporting Information Figure S6).

An alternative classification of plastic responses, independent of reversion/reinforcement, is to assess whether the plastic response of Control-lines brings expression levels closer to the new optimum, i.e. expression levels of native selection-lines in the treatment (*La* in Figure 1). Such a pattern can occur even if plastic responses of Control-lines and evolutionary changes are in opposite directions and classified as reversion (for example Figure 1C, reversion with 'overshooting'), but may indicate that ancestral plasticity was beneficial. We therefore tested for all genes with considerable plastic changes (> 0.2 *Lo*) whether expression levels of adapted selection lines (*La*) in the treatment were closer to the plastic levels of Control-lines (*Lp*) or whether they showed a compensation of the plastic response and were more similar to the ancestral stage (*Lo,* Control-lines in control conditions). Such a situation would strongly indicate that plastic responses were maladaptive and expression levels of Control-lines in the treatment not beneficial in the long-term.

To better understand the relationship between the within-line proportions of reversed genes and proportions of genes with *La* closer to *Lo* with adaptation, we calculated the Spearman correlation between the proportions of reversed expression changes (or *La* closer to *Lo* respectively) and mean offspring number in seven selection lines in HD. We focused on HD because it was the most extreme environment with the strongest decline in offspring number. To test for significance, we used permutations: Mean offspring numbers were randomly assigned to lines and correlation was calculated again. Plastic and evolutionary changes were defined using a cutoff of 20 % *Lo.* Proportion of permutations with a correlation coefficient exceeding the observed value gave the respective *P*-value.

## Results

*Fitness assay showed evolutionary adaptation*

We found that selection lines had a higher offspring number in their native condition compared to Control-lines (Dry: $F_{1,14} = 9.20$, $p = 0.009$; Hot: $F_{1,16} = 4.78$, $p = 0.044$; Hot-Dry: $F_{1,16} = 23.51$, $p = 1.786E-04$), confirming that adaptation had occurred (Figure 2). In contrast to treatment conditions, there was no difference in offspring number between Control- and selection-lines under control conditions (CT-lines: 64.89 [61.49, 68.29]; D-lines: 65.76 [61.26, 70.25], H-lines: 67.41 [63.76, 71.06], HD-lines: 65.56 [62.17, 68.94]) (Figure 2). Using three additional mixed models, we compared how lines from different selection regimes responded to treatments. We found significant negative effects for all stress treatments (D: $F_{1,23} = 45.37$, $p = 6.85E-07$; H: $F_{1,28} = 507.68$, $p < 2.200E-16$; HD: $F_{1,28} = 553.06$, $p < 2.200E-16$) (Figure 2). Interaction between selection and treatments, i.e. whether the response to the treatment was different

depending on selection regime, was significant for HD-lines ($F_{1,28}$ = 9.39, $p$= 4.754E-03) and for D ($F_{1,23}$ = 4.32, $p$ = 0.049), but not for H ($F_{1,28}$ = 0.57, $p$ = 0.455). Interaction between selection regime, treatment and line type (normal/mixed) was not significant in any treatment (see Table S3), indicating that mixing lines did not have an effect on how they respond to the treatment. However, focusing on offspring number within each treatment separately we obtained significant results for mixing in Hot-Dry with the mixed lines having higher offspring numbers than normal lines (see also Figure S2). ANOVA tables as well as results of the linear mixed models are in Supporting information Table S3.

*Plastic responses are stronger than evolutionary changes and responsible for most of the total divergence*

We found that the treatments induced significant plastic as well as evolutionary changes in gene expression (Table 2). Plastic responses of Control-lines induced by the combined stress treatment Hot-Dry included the highest number of DE genes (4651), Dry resulted in the smallest number of responding genes (365, Table 2). The same was true for evolutionary changes, i.e. differences between Control-lines and selection-lines in the treatment. Only 18 genes showed significant differences in Dry, 25 in Hot and 55 in Hot-Dry. Total divergence (difference Control-lines in control and selection-lines in treatment, Figure 1, Figure S1) was highest between Control and Hot-Dry (2045 genes) and lowest between Control and Dry (283). Therefore, of those genes with ancestral plasticity, only a small proportion exhibited evolutionary responses in the treatments (Supporting Information Figure S4). Quantifying the relative contribution of evolutionary change and ancestral plasticity to the observed total 2-fold expression divergence (TC in Figure 1), we found larger contribution of ancestral plasticity than of evolutionary changes in Hot-Dry and Hot (median proportions: Hot: 71.47%; Hot-Dry: 86.30%), but not in Dry (42.90%). Many genes had ancestral plastic responses with magnitudes higher than the observed total divergence especially in Hot-Dry. In these cases, the evolutionary change was in opposite direction to ancestral plasticity and (partly) reversed it. Total divergence was thus reduced to the point that about half of the DE genes with evolutionary changes in Hot and Hot-Dry did not diverge significantly from their ancestral expression levels in Control (see also Figure S4).

*Evolutionary responses are more likely among plastic genes and mainly opposed to plastic responses*

Understanding how ancestral plasticity and evolutionary changes interacted, we found that plastic genes were overrepresented among evolved genes (Dry: P = 8.77E-04 Fisher's exact test; Hot: P = 0.026; Hot-Dry: P = 0.043) suggesting that plasticity does not prevent subsequent evolution. (Please note that

Control-lines were split into two groups to obtain independent estimates of plastic levels in treatments, see Methods). In almost all genes with significant evolutionary changes and ancestral plastic responses, the two responses were in opposite direction (Dry: 2 out of 2 genes; Hot: 12 out of 13; Hot-Dry: 33 out of 34).

DE analysis is designed to keep the number of false positives low, which is important for identifying candidate genes in a highly dimensional data set. However, we may lose some information by focusing only on those highly significant gene set. Since our primary aim here was not to identify a limited number of candidate genes, but instead to get a more comprehensive overview of evolutionary and plastic responses, we additionally used mean expression levels per selection line (see Methods) following an approach proposed by Ho and Zhang (2018). We calculated for each gene the mean of Control-lines in control (original level *Lo*), the adapted level (*La*), which is the mean of a selection-line in the treatment and the plastic level (*Lp*), expression of control lines in the treatment. Control-lines were split into two groups to obtain two independent estimates of *Lp*. From these measurements we can test whether the plastic response (difference between *Lp* and *Lo*) and evolutionary responses (difference between *Lp* and *La*) are in the same direction. We obtained results consistent with the DE analysis: Reinforcements were less frequent than reversions in three of the four selection lines in Dry, in three of five Hot-Lines and in all HD lines (binomial test $P < 0.05$), see Figure 5 (Figure S5 A-C for selection-lines separately). Parametric bootstrapping (Ho & Zhang, 2019) confirmed our results (Supporting information, Figure S5 D-E). In Dry three out of four lines showed a significant prevalence of reversions over reinforcements ($P < 0.05$ in binomial test, Supporting Information Figure S5 D), in Hot and in Hot-Dry reversions were more frequent in all lines (Figure S5 E and F). In many of these genes, expression levels returned to original levels *Lo* or showed an overcompensation below *Lo* (Dry: 58.82 ± 1.7 %; Hot: 35.72 ± 7.8 %; Hot-Dry: 37.56 ± 1.5 %). However, for quantifying reversions, only those genes showing substantial evolutionary changes were included, i.e. we excluded plastic genes with expression levels close (difference smaller than cut-off of 20 % Control-levels) to levels of adapted lines in the treatment ('overshooting' in Figure 1 C). To include these plastic responses that were potentially close to the adapted level in treatment conditions, we tested for all genes with large ancestral plasticity whether adapted levels in selection-lines *(La)* were closer to their ancestral plastic levels (*Lp*) or to original levels in control conditions (*Lo*). We found that for a large proportion of ancestrally plastic genes, the evolved level of adapted selection-lines (*La*) was closer to their ancestral expression in Control (*Lo*) than to the plastic response level in treatment (*Lp*) (Dry: 45.4 ± 1.1%; Hot: 41.2 ± 1.2%; Hot-Dry: 40.7 ± 2.5 %) (Figure 5B) suggesting that plastic levels of Control-lines were not beneficial for long-term adaptation.

*Evolution of plasticity: Little change in plasticity in selection lines*

Number of genes with significant changes in plasticity as detected in the *limma* analysis was small in all conditions (only five genes in total, Table 2). As mentioned previously, stringent p-value adjustment can severely limit our power to detect subtle changes in the degree of plasticity. Higher numbers of plastic genes in Control-Lines compared to selection-lines (in Dry: 365 vs 49; Hot: 3479 vs 2798; Hot-Dry: 4651 vs 3119, Table 2) suggested that adapted lines lost plasticity in some genes. These differences were significant in Hot-Dry (based on a permutation test Table 3, see Methods). We also found that the magnitude of plastic changes was significantly smaller in adapted selection-lines in Dry and Hot-Dry (Table 3, see also Figure 4).

We were then interested to examine whether the evolutionary changes in the treatments ($EC_T$ in Figure 1) were due to a change in mean expression (shift in the intercept of reaction norms, see Figure 1D) or to a change in plasticity (different slopes of reaction norms, see Figure 1E). We then quantified the relative contributions of changes in the mean versus changes in plasticity to the observed evolutionary change in the treatment ($EC_T$ in Figure 1), and found that evolution of the intercept explained more evolutionary divergence than evolution of the slope of reaction norms, especially in Dry (Mann-Whitney U test: Dry: U = 312, *P* = 5.994e-08; Hot: U = 136, P = 4.34E-04; Hot-Dry: U = 1039, *P* = 0.012;) (Figure 3). A shift in the intercept, i.e. in the overall mean, should not only lead to differences in treatment conditions but also result in a correlated change in control conditions (see scenario in Figure 1D, $EC_{CT}$ and $EC_T$). We found indeed that differences between selection- and Control-lines in control and treatment conditions were positively correlated ($P < 0.001$, permutation test, Table 3).

*Proportions of reversed plastic responses and association with fitness*

To gain a better understanding of the adaptive value of the changes of expression levels in the evolved lines, we tested for an association between within-line proportion of reversed or reinforced plastic responses and the average fitness of the lines in the HD treatment. We found that lines with a higher proportion of reversions had a higher average offspring number (correlation: 0.82, *P* = 0.012, Figure 6A) and we found a negative but non-significant correlation between fitness and reinforcements (correlation: -0.43, *P* = 0.85). When we tested for an association between fitness and proportion of ancestrally plastic genes with *La* closer to *Lo*, we also found a positive correlation (correlation: 0.86, *P* = 0.006, Figure 6B). Overall, better adapted lines (higher fitness in HD) showed a higher proportion of reversed ancestral

plasticity and these plastic genes were more similar to the original expression levels of control lines in CT. Performing the analysis with gene expression data in H and D provided similar correlations, although not significant because of lower sample sizes (Supporting information, Figure S8, Table S4.2).

## Discussion

We studied plastic and evolved responses in gene expression of *T. castaneum* in response to three new environmental conditions (Dry, Hot, Hot-Dry). After 20 generations of experimental evolution, we were able to detect adaptation and found significant evolutionary changes in expression levels. Comparing evolutionary changes with ancestral plastic responses showed that a reversion of plasticity was most frequent (> 95 % genes with significant plastic changes in DE analysis; > 40 % of genes with substantial plastic changes, i.e. changes higher than 20% of ancestral levels). The number of genes where ancestral plasticity was reinforced by evolution was significantly smaller (DE analysis: <5%, 27-34 % of genes with substantial plastic changes). A high proportion of the originally plastic genes evolved to expression levels that were closer to control levels than to ancestrally plastic levels. Although the proportion of non-reversed plastic genes was still high, positive associations between fitness and proportion of reversions, and compensated plasticity (expression levels closer to control levels) respectively, suggest that ancestral plasticity was maladaptive for a majority of responding genes. Although plasticity showed a high degree of preservation in terms of number of responding genes and direction of the response, we found evidence that selection lines evolved a reduced plasticity and thus partly compensated the maladaptive ancestral response. We were further able to show a positive association between the proportion of reversed plastic responses and adaptation (mean fitness per line) in the most stressful treatment Hot-Dry. We found that the effect of increased temperature was much stronger than reduced humidity. A strong response to heat is expected in ectothermic organisms, whereas humidity reduction is probably less stressful for *Tribolium* due to specific adaptations to this environment (Park & Beeman, 2008; Sokoloff, 1972) (see also results of a previous study: Koch and Guillaume (2020) ).

*Reversion of plastic responses*

Different patterns describing the relationship between plastic and evolved changes in gene expression have been documented. It was suggested that plasticity might help populations to persist after environmental change, or to colonize new habitats by bringing phenotypes closer to the new optimum. Studies found support for this hypothesis by showing that plastic responses of non-adapted individuals diminished differences to native populations (Lohman, Stutz, & Bolnick, 2017; Mäkinen, Papakostas,

Vøllestad, Leder, & Primmer, 2016). Adaptive plasticity can also be indicated when plastic and evolutionary responses are in the same direction (Li, Li, Song, Wang, & Zhang, 2017), or when plasticity is higher in adapted populations (Hasan et al., 2017; McCairns & Bernatchez, 2009), suggesting that most plastic individuals were favored by selection. However, there are also examples for the reversed pattern suggesting that plasticity was maladaptive. In wild populations of *Fundulus heteroclitus*, evolved changes to different temperatures were opposite to plastic responses of the ancestral population (Dayan et al., 2015). *Rhagoletis* flies shifting to a new host fruits showed evolutionary responses opposite to plasticity of non-adapted species (Ragland et al., 2015). Experimental evolution studies found countergradient evolution in *Drosophila* adapting to different diets (Huang & Agrawal, 2016; Yampolsky et al., 2012) and in guppies adapting to low predation environments (Ghalambor et al., 2015). A comparative study (Ho & Zhang, 2018) analyzing data of multiple experimental evolution suggested that reversions of gene expression changes might be a general pattern during adaptation.

Our study fits with these previous observations. We found a higher proportion of reversions than reinforcements in all conditions indicating mostly maladaptive plasticity. An alternative explanation for the prevalence of reversions without maladaptive plastic responses would be that control lines exhibited a response in the right direction, but overshot an optimum expression level (Figure 1C). Fine-tuning during long-term adaptation could then lead to a partial reversion of the plastic response. We took this possibility into account by not only focusing on a reversion of plasticity, but also testing whether plastic changes brought expression levels closer to the adapted level (*La*) than to the ancestral level of expression in Control (*Lo*). If, in contrast, we see that the adapted level of expression (*La*) is closer to the original level (*Lo*, CT lines in CT conditions), it indicates that plastic responses were maladaptive since they moved expression levels further away from the new optimum and became compensated during evolution (see Figure 1C). We found that adapted lines showed a high proportion of expression levels closer to their ancestral level (*Lo*) than to their ancestral plastic response (*Lp*). We could further show that this proportion is positively associated with higher fitness per selection line. We also found strong positive associations of within-line proportions of reversions with mean reproductive output, indicating a possible fitness advantage to reversions.

There was still a large proportion of plastic genes that did not show reversion. They are either close to the levels of adapted lines or showed reinforcement. The first case might indicate that plasticity prevented evolution by matching the new optimum. Reinforced plastic changes could be examples of adaptive plasticity. However, correlation between proportion of reinforcements and fitness was not significant and negative. The observed positive correlation between proportion of reversed plastic genes and mean fitness per line in HD rather suggests that reversions were favoured during adaptations. Reversions might

become more pronounced after more generations, once evolution had sufficient time to further reverse maladaptive ancestral plastic responses. Our selection lines still show a strong reduction in offspring number compared to control levels suggesting further potential to adapt.

*Evolution of reaction norms*

The ancestral maladaptive plasticity can be compensated by shifts in the intercept or changes in the slope of reaction norms. Both are not mutually exclusive and can occur together in the same trait. We were aiming to quantify their relative importance for evolutionary responses in transcriptomes.

Gene expression studies so far provided mixed results regarding the evolution of plasticity. *Drosophila* populations adapted to different temperatures showed local adaptation, but there was no evidence for evolution of thermal reaction norms of different transcripts and changes affected mainly expression mean (Clemson, Sgrò, & Telonis-Scott, 2016). Experimental evolution studying *Drosophila* under variable diets found no significant changes in plasticity (Yampolsky et al., 2012) or less than expected (Huang & Agrawal, 2016). In contrast, other studies found differences in temperature responses between tropical and temperate *Drosophila* populations (Levine, Eckert, & Begun, 2011; von Heckel et al., 2016). Other examples for differences in genes expression plasticity between adapted and non-adapted populations include temperature (Morris et al., 2014) and salinity (Gibbons, Metzger, Healy, & Schulte, 2017; McCairns & Bernatchez, 2009) responses of marine and freshwater sticklebacks, temperature response of killifish populations from different latitudes, as well as plastic responses to toxic hydrogen sulphide ($H_2S$) of fish population from $H_2S$ rich springs versus non-toxic springs (Passow et al., 2017). There was no consistent pattern regarding the direction in which plasticity evolves: In some cases adapted population showed an increase in plasticity (Morris et al., 2014), in other cases plasticity was reduced (Huang & Agrawal, 2016; Ragland et al., 2015; von Heckel et al., 2016) or reduction and enhancement of plasticity were equally frequent (Gibbons et al., 2017; Yampolsky et al., 2014). Overall, there is evidence in multiple species that expression plasticity of some genes can evolve. However, even in some of these studies reporting evolved plasticity (Dayan et al., 2015; Gibbons et al., 2017; Morris et al., 2014) the number of transcripts with significant changes in the mean was much higher than transcripts with changed plasticity and large parts of the plastic responses showed a high degree of preservation.

In accordance with these previous findings we found that changes in the mean contributed more to the observed expression differences in the treatments than changes in plasticity. A possible reason might be that genetic variation in mean expression was higher than genetic variation in plasticity. In addition, we

did not select for changes in plasticity directly since the conditions in the treatments were constant. Selection was therefore on expression levels in the treatment and only indirectly on plasticity. Plasticity could evolve if mean expression levels were genetically correlated with plasticity. Although plastic responses showed a high degree of preservation in terms of affected genes and direction, we still found some evidence for evolutionary changes in the magnitude of plastic responses, i.e. the slope of the reaction norm (Via, 1993).

*Why should ancestral plastic responses be reversed?*

New stressors might disturb homeostasis resulting in inappropriate responses, and long-term adaptation therefore restores ancestral phenotypes by genetic changes, referred to as genetic compensation (Grether, 2005) or counter-gradient variation (Conover et al., 2009). However, in our study we applied relatively mild stressor treatments, i.e. individuals were able to survive and reproduce. Drought and heat are also stressors, which *T. castaneum* had experienced in the past (Sokoloff, 1972), so there had probably been selection on plastic responses to be beneficial. However, plastic responses might be optimized for a short-term exposure: Allocation of resources from reproduction to protection might increase survival probability and allow individuals to continue reproduction as soon as the stress has disappeared, but this response becomes maladaptive during continuous exposure and should therefore be under negative selection. Expression of stress related genes is in general accompanied by a down-regulation of genes involved in growth and reproduction due to an allocation of resources (Schwenke, Lazzaro, & Wolfner, 2016; Sokolova, 2013). A well-studied example are heat shock proteins (hsp). Hsp are well known for their protective function and to be crucial for survival (Feder & Hoffman, 1999), but it was also shown that their expression comes at a cost (Feder et al., 1998; Sørensen, Kristensen, & Loeschcke, 2003). Accordingly, it was often found that hsp expression in populations adapted to warmer climates is lower compared to non-adapted populations (Fangue, Hofmeister, & Schulte, 2006; Narum & Campbell, 2015; J G Sørensen, Dahlgaard, & Loeschcke, 2001). In general, other protection mechanisms independent of ancestral plasticity may arise during long-term adaptation (e.g. enzymes, which are more stable at high temperature) and make the costly stress response expendable.

An alternative explanation for the reduced plasticity in adapted lines is that the signal responsible for eliciting the plastic responses is based on any kind of damage (e.g. deformations in macromolecules, membrane lipids, proteins, and DNA) caused by heat or stress in general (Kültz, 2005). Higher resistance in adapted lines might shift the inducing thresholds, i.e. the temperature when damages occur and stress

response is induced (Sikkink, Reynolds, Ituarte, Cresko, & Phillips, 2014) above the levels we applied in the treatments.

Interestingly, we found no differences in fitness between lines from different selection regimes under Control conditions (Figure 2). We could detect some genes that differed in expression levels between Control-lines and selection-lines in control conditions, but this did not seem to affect offspring number. It indicates a lack of fitness trade-offs, where alleles providing a fitness advantage in one environment (treatment) are detrimental in another (Control). Together with the observation that selection lines evolved to bring expression closer to ancestral expression levels, it suggests that for many genes the optimal expression level is not different between conditions. They might be involved in processes important for maintenance and reproduction. Under stress, limited resources have to be invested into protection, that are then not available for reproduction (Sokolova, 2013). Long-term adaptation should then work to restore control levels that are likely to be optimized for highest reproductive output and to reduce costly stress responses resulting in improved canalization of traits associated with fitness (Stearns & Kawecki, 1994). Canalization, i.e. robustness against environmental variation, was found previously in gene expression adaptation (Levine et al., 2011; Shaw et al., 2014; von Heckel et al., 2016). Genetic differences between control and selection lines that are responsible for adaptation to the treatments did not have an effect in control conditions and thus represent cryptic genetic variation (Gibson & Dworkin, 2004). They might either concern genes that are not expressed in control conditions or represent changes neutral under control conditions.

*Potential caveats*

The number of genes with significant plastic changes in the DE analysis was much higher compared to genes showing evolutionary changes. One possible explanation would be that adaptive plasticity prevented evolution. If the plastic responses matched the optimum, no genetic changes in the selection lines are expected to occur. However, when we analysed each line separately and considered a gene as evolved if the mean difference between its adapted expression level (*La*) and ancestral plastic expression level (*Lp*) was more than 20% of the ancestral expression level in Control (*Lo*), we found approximately the same number of genes with evolutionary change and ancestral plasticity (Supporting information Table S4.1).

In the DE analysis in *limma* we did not analyse each line separately but treated them as biological replicates. Since lines were split across conditions, comparisons between conditions, i.e. plastic changes,

can be made within lines. They should thus be more precise and statistical power should be higher than comparisons between selection regimes, i.e. evolved changes, that have to be made between lines. Differences between lines from the same selection regime lower the ability to obtain significant evolutionary changes. These differences can arise from genetic drift. Since our population size was relatively small (120 individuals per line) this might have been an important factor. Another explanation is that lines from the same selection regime differed how exactly they improved their fitness in the respective treatment. Since fitness is a highly polygenic trait, the genes contributing to a fitness increase may not be the same in different lines (see Barghi et al., 2019). For the most extreme treatment HD, where we sequenced seven lines, we further found considerable differences in fitness between the lines, suggesting that not all of them were at the same stage of adaptation. It is therefore not surprising that expression levels did not evolve in the same way among lines.

The DE analysis in *limma* requires that a gene shows similar changes in all replicate lines and is therefore more conservative. If the main interest of a study is to identify promising candidate genes for future more detailed analyses, it is the appropriate approach to keep FDR as low as possible. In contrast, if the focus is more on general patterns, a less stringent analysis using mean expression levels can give us a more complete picture. Since genetic drift is random it cannot explain the observed excess of reversions over reinforcements.

Although we demonstrated that gene expression changed during evolution, it is not clear whether these changes are the cause of an increased fitness in these conditions or whether they are rather the consequence of adaptation and being less stressed. One disadvantage in studying whole transcriptomes is that not all responding genes might be of functional importance but are correlated to other adaptive changes. High intercorrelations within the transcriptome (Ayroles et al., 2009; McGraw et al., 2011) might lead to correlated responses in many other genes. Furthermore, observed evolutionary changes might be caused by indirect selection and other mechanisms, e.g. changes in protein structure of enzymes, were responsible for adaptation of selection lines. Future studies that manipulate expression and test for correlated changes in offspring are needed to confirm adaptive value of expression changes.

**Conclusions**

We found that genes with the strongest plastic responses showed evolutionary changes in opposite directions, suggesting that ancestral plasticity was maladaptive for long-term adaptation. In the most stressful treatment (Hot-Dry), selection lines with higher fitness show a higher proportion of reversions and a higher proportion of originally plastic genes that are closer to ancestral expression levels.

Differences between adapted lines and control lines in the treatment were mainly due to a change in mean expression (i.e. shift in the intercept of reaction norms), while plasticity was preserved in terms of affected genes and direction of change. However, we found that a part of the differences in the treatments can be explained by a reduction in the magnitude of plasticity in adapted lines. Our results add to growing evidence that plasticity and evolution are often in opposite direction and maladaptive plastic responses might increase the strength of selection. In contrast to previous studies, we included fitness data in our analyses, which allowed us to give evidence for adaptation. Furthermore, we were able to show an association between reversion of plasticity and adaptation in the most stressful condition. Similar results in all three stress treatments indicated that these findings may represent a general pattern of gene expression adaptation.

**Data Accessibility**

RNA-seq data and read counts for measuring gene expression are available from GEO, accession number GSE156256. Fitness data and code for analysing fitness and gene expression are available at Zenodo https://doi.org/10.5281/zenodo.3980629.

**Author Contribution**

FG and ELK designed experiment. ELK conducted experiment, laboratory work and analysed the data. FG and ELK wrote the manuscript.

**Acknowledgements**

This work was supported by the Swiss National Science Foundation, grants PP00P3_1144846 and PP00P3_176965 to FG. We thank Sonja Sbilordo, Sara Meier, Cilgia Lippuner, Tim Schoch and Tim Emmenegger for helping during the fitness assay, Elke Karaus, Martina Berchtold, Encarnación Lozano for RNA extraction and quality control, Lucy Poveda and Maria Domenica Moccia for advice in library preparation and sequencing, Lennart Opitz for helping to process the RNA-seq data.

## References


Alvarez, M., Schrey, A. W., & Richards, C. L. (2015). Ten years of transcriptomics in wild populations: What have we learned about their ecology and evolution? *Molecular Ecology*, *24*(4), 710–725. doi:10.1111/mec.13055

Ayroles, J. F., Carbone, M. A., Stone, E. A., Jordan, K. W., Lyman, R. F., Magwire, M. M., … Mackay, T. F. C. (2009). Systems genetics of complex traits in *Drosophila melanogaster*. *Nature Genetics*, *41*(3), 299–307. doi:10.1038/ng.332

Barghi, N., Tobler, R., Nolte, V., Jakšić, A. M., Mallard, F., Otte, K. A., … Schlötterer, C. (2019). *Genetic redundancy fuels polygenic adaptation in Drosophila*. *PLoS Biology* (Vol. 17). doi:10.1371/journal.pbio.3000128

Barton, M., Sunnucks, P., Norgate, M., Murray, N., & Kearney, M. (2014). Co-gradient variation in growth rate and development time of a broadly distributed butterfly. *PLoS ONE*, *9*(4), 1–8. doi:10.1371/journal.pone.0095258

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

Carroll, S. B. (2005). Evolution at Two Levels: On Genes and Form. *PLoS Biology*, *3*(7), e245. doi:10.1371/journal.pbio.0030245

Chevin, L. M., Lande, R., & Mace, G. M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLoS Biology*, *8*(4). doi:10.1371/journal.pbio.1000357

Clemson, A. S., Sgrò, C. M., & Telonis-Scott, M. (2016). Thermal plasticity in Drosophila melanogaster populations from eastern Australia: quantitative traits to transcripts. *Journal of Evolutionary Biology*, *29*(12), 2447–2463. doi:10.1111/jeb.12969

Conover, D. O., Duffy, T. A., & Hice, L. A. (2009). The covariance between genetic and environmental influences across ecological gradients: Reassessing the evolutionary significance of countergradient and cogradient variation. *Annals of the New York Academy of Sciences*, *1168*(1), 100–129.

doi:10.1111/j.1749-6632.2009.04575.x

Crispo, E. (2007). The Baldwin effect and genetic assimilation: Revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution*, *61*(11), 2469–2479. doi:10.1111/j.1558-5646.2007.00203.x

Dayan, D. I., Crawford, D. L., & Oleksiak, M. F. (2015). Phenotypic plasticity in gene expression contributes to divergence of locally adapted populations of *Fundulus heteroclitus*. *Molecular Ecology*, *24*(13), 3345–3359. doi:10.1111/mec.13188

de Jong, G. (2005). Evolution of phenotypic plasticity: patterns of plasticity and the emergence of ecotypes. *New Phytologist*, *166*(1), 101–118. doi:10.1111/j.1469-8137.2005.01322.x

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21.

Draghi, J. A., & Whitlock, M. C. (2012). Phenotypic plasticity facilitates mutational variance, genetic variance, and evolvability along the major axis of environmental variation. *Evolution*, *66*(9), 2891–2902. doi:10.1111/j.1558-5646.2012.01649.x

Fangue, N. A., Hofmeister, M., & Schulte, P. M. (2006). Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. *The Journal of Experimental Biology*, *209*(Pt 15), 2859–72. doi:10.1242/jeb.02260

Feder, M. E., & Hoffman, G. E. (1999). Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annual Review of Physiology*, *61*, 243–282.

Feder, M. E., & Krebs, R. A. (1998). Natural and Genetic Engineering of the Heat-Shock Protein Hsp70 in Drosophila melanogaster : Consequences for Thermotolerance. *American Zoologist*, *38*(3), 503–517. doi:10.1093/icb/38.3.503

Fierst, J. L. (2011). A history of phenotypic plasticity accelerates adaptation to a new environment. *Journal of Evolutionary Biology*, *24*(9), 1992–2001. doi:10.1111/j.1420-9101.2011.02333.x

Fitzpatrick, B. M. (2012). Underappreciated consequences of phenotypic plasticity for ecological speciation. *International Journal of Ecology*, *2012*, 32–37. doi:10.1155/2012/256017

Forsman, A. (2015). Rethinking phenotypic plasticity and its consequences for individuals, populations and

669 species. *Heredity*, *115*(4), 276–284. doi:10.1038/hdy.2014.92

670 Garland, T., & Kelly, S. A. (2008). Phenotypic plasticity and experimental evolution. *Journal of*
671 *Experimental Biology*, *211*(16), 2725–2725. doi:10.1242/jeb.022673

672 Gavrilets, S., & Scheiner, S. M. (1993). The genetics of phenotypic plasticity. V. Evolution of reaction norm
673 shape. *Journal of Evolutionary Biology*, *6*(1), 31–48. doi:10.1046/j.1420-9101.1993.6010031.x

674 Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N., & Hughes, K. A. (2015). Non-
675 adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature*,
676 *525*(7569), 372–375. doi:10.1038/nature15256

677 Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive
678 phenotypic plasticity and the potential for contemporary adaptation in new environments.
679 *Functional Ecology*, *21*(3), 394–407. doi:10.1111/j.1365-2435.2007.01283.x

680 Gibbons, T. C., Metzger, D. C. H., Healy, T. M., & Schulte, P. M. (2017). Gene expression plasticity in
681 response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats.
682 *Molecular Ecology*, *26*(10), 2711–2725. doi:10.1111/mec.14065

683 Gibson, G. (2008). The environmental contribution to gene expression profiles. *Nature Reviews Genetics*,
684 *9*(8), 575.

685 Gibson, G., & Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics*, *5*(9), 681–
686 690. doi:10.1038/nrg1426

687 Grether, G. F. (2005). Environmental Change, Phenotypic Plasticity, and Genetic Compensation. *The*
688 *American Naturalist*, *166*(4), E115–E123. doi:10.1086/432023

689 Guo, J., Liu, R., Huang, L., Zheng, X. M., Liu, P. L., Du, Y. S., … Ge, S. (2016). Widespread and adaptive
690 alterations in genome-wide gene expression associated with ecological divergence of two oryza
691 species. *Molecular Biology and Evolution*, *33*(1), 62–78. doi:10.1093/molbev/msv196

692 Hasan, M. M., DeFaveri, J., Kuure, S., Dash, S. N., Lehtonen, S., Merilä, J., & McCairns, R. J. S. (2017).
693 Sticklebacks adapted to divergent osmotic environments show differences in plasticity for kidney
694 morphology and candidate gene expression. *The Journal of Experimental Biology*, *220*(12), 2175–
695 2186. doi:10.1242/jeb.146027

Hatakeyama, M., Opitz, L., Russo, G., Qi, W., Schlapbach, R., & Rehrauer, H. (2016). SUSHI: An exquisite recipe for fully documented, reproducible and reusable NGS data analysis. *BMC Bioinformatics*, *17*(1), 1–9. doi:10.1186/s12859-016-1104-8

Ho, W.-C., & Zhang, J. (2018). Evolutionary adaptations to new environments generally reverse plastic phenotypic changes. *Nature Communications*, *9*(1), 1–11. doi:10.1038/s41467-017-02724-5

Ho, W.-C., & Zhang, J. (2019). Genetic gene expression changes during environmental adaptations tend to reverse plastic changes even after the correction for statistical nonindependence. *Molecular Biology and Evolution*. doi:10.1093/molbev/msz002

Hodgins-Davis, A., & Townsend, J. P. (2009). Evolving gene expression: from G to E to G × E. *Trends in Ecology and Evolution*, *24*(12), 649–658. doi:10.1016/j.tree.2009.06.011

Huang, Y., & Agrawal, A. F. (2016). Experimental Evolution of Gene Expression and Plasticity in Alternative Selective Regimes. *PLoS Genetics*, *12*(9), 1–23. doi:10.1371/journal.pgen.1006336

Koch, E. L., & Guillaume, F. (2020). Additive and mostly adaptive plastic responses of gene expression to multiple stress in Tribolium castaneum. *PLoS Genetics*, *16*(5), e1008768. doi:10.1371/journal.pgen.1008768

Krebs, R. A., & Feder, M. E. (1997). Natural Variation in the Expression of the Heat-Shock Protein HSP70 in a Population of Drosophila melanogaster and Its Correlation with Tolerance of Ecologically Relevant Thermal Stress. *Evolution*, *51*(1), 173. doi:10.2307/2410970

Kregel, K. C. (2002). Invited Review: Heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. *Journal of Applied Physiology*, *92*(5), 2177–2186. doi:10.1152/japplphysiol.01267.2001

Kültz, D. (2005). Molecular and Evolutionary Basis of the Cellular Stress Response. *Annual Review of Physiology*, *67*(1), 225–257. doi:10.1146/annurev.physiol.67.040403.103635

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. doi:10.18637/jss.v082.i13

Lande, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology*, *22*(7), 1435–1446. doi:10.1111/j.1420-9101.2009.01754.x

Laugen, A. T., Laurila, A., Räsänen, K., & Merilä, J. (2003). Latitudinal countergradient variation in the common frog ( Rana temporaria ) development rates – evidence for local adaptation. *J. Evol.*, *16*(5), 996–1005.

Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology*, *15*(2), R29. doi:10.1186/gb-2014-15-2-r29

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. doi:10.18637/jss.v069.i01

Levine, M. T., Eckert, M. L., & Begun, D. J. (2011). Whole-genome expression plasticity across tropical and temperate *Drosophila melanogaster* populations from eastern Australia. *Molecular Biology and Evolution*, *28*(1), 249–256. doi:10.1093/molbev/msq197

Levis, N. A., & Pfennig, D. W. (2016). Evaluating ' Plasticity-First ' Evolution in Nature : Key Criteria and Empirical Approaches. *Trends in Ecology & Evolution*, *31*(7), 563–574. doi:10.1016/j.tree.2016.03.012

Li, A., Li, L., Song, K., Wang, W., & Zhang, G. (2017). Temperature, energy metabolism, and adaptive divergence in two oyster subspecies. *Ecology and Evolution*, *7*(16), 6151–6162. doi:10.1002/ece3.3085

Liao, Y., Smyth, G. K., & Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. doi:10.1093/bioinformatics/btt656

Loeschcke, V., & Hoffmann, A. a. (2007). Consequences of heat hardening on a field fitness component in Drosophila depend on environmental temperature. *American Naturalist*, *169*(2), 175–183. doi:10.1086/510632

Lohman, B. K., Stutz, W. E., & Bolnick, D. I. (2017). Gene expression stasis and plasticity following migration into a foreign environment. *Molecular Ecology*, *26*(18), 4657–4670. doi:10.1111/mec.14234

Mahroof, R., Zhu, K. Y., & Subramanyam, B. (2005). Changes in Expression of Heat Shock Proteins in Tribolium castaneum (Coleoptera: Tenebrionidae) in Relation to Developmental Stage, Exposure Time, and Temperature. *Annals of the Entomological Society of America*, *98*(1), 100–107.

752 doi:10.1603/0013-8746(2005)098[0100:CIEOHS]2.0.CO;2

753 Mäkinen, H., Papakostas, S., Vøllestad, L. A., Leder, E. H., & Primmer, C. R. (2016). Plastic and evolutionary
754 gene expression responses are correlated in European grayling (*Thymallus thymallus*)
755 subpopulations adapted to different thermal environments. In *Journal of Heredity* (Vol. 107, pp. 82–
756 89). doi:10.1093/jhered/esv069

757 Mallard, F., Jakšić, A. M., & Schlötterer, C. (2018). Contesting the evidence for non-adaptive plasticity.
758 *Nature*, *555*(7698), E21.

759 McCairns, R. J. S. S., & Bernatchez, L. (2009). Adaptive divergence between freshwater and marine
760 sticklebacks: insights into the role of phenotypic plasticity from an integrated analysis of candidate
761 gene expression. *Evolution*, *64*(4), 1029–1047. doi:10.1111/j.1558-5646.2009.00886.x

762 McGraw, E. A., Ye, Y. H., Foley, B., Chenoweth, S. F., Higgie, M., Hine, E., & Blows, M. W. (2011). High-
763 dimensional variance partitioning reveals the modular genetic basis of adaptive divergence in gene
764 expression during reproductive character displacement. *Evolution*, *65*(11), 3126–3137.
765 doi:10.1111/j.1558-5646.2011.01371.x

766 Milutinović, B., Stolpe, C., Peuß, R., Armitage, S. A. O., & Kurtz, J. (2013). The Red Flour Beetle as a Model
767 for Bacterial Oral Infections. *PLoS ONE*, *8*(5). doi:10.1371/journal.pone.0064638

768 Molina-Montenegro, M. A., Peñuelas, J., Munné-Bosch, S., & Sardans, J. (2012). Higher plasticity in
769 ecophysiological traits enhances the performance and invasion success of Taraxacum officinale
770 (dandelion) in alpine environments. *Biological Invasions*, *14*(1), 21–33. doi:10.1007/s10530-011-
771 0055-2

772 Morris, M. R. J., Richard, R., Leder, E. H., Barrett, R. D. H., Aubin-Horth, N., & Rogers, S. M. (2014). Gene
773 expression plasticity evolves in response to colonization of freshwater lakes in threespine
774 stickleback. *Molecular Ecology*, *23*(13), 3226–3240. doi:10.1111/mec.12820

775 Narum, S. R., & Campbell, N. R. (2015). Transcriptomic response to heat stress among ecologically
776 divergent populations of redband trout. *BMC Genomics*, *16*(1), 103. doi:10.1186/s12864-015-1246-5

777 Nussey, D. H., Postma, E., Gienapp, P., & Visser, M. E. (2005). Selection on heritable phenotypic plasticity
778 in a wild bird population. *Science*, *310*(5746), 304–306. doi:10.1126/science.1117004

779 Park, Y., & Beeman, R. W. (2008). Postgenomics of Tribolium: Targeting the endocrine regulation of

diuresis. *Entomological Research*, *38*(2), 93–100. doi:10.1111/j.1748-5967.2008.00143.x

Passow, C. N., Henpita, C., Shaw, J. H., Quackenbush, C. R., Warren, W. C., Schartl, M., … Tobler, M. (2017). The roles of plasticity and evolutionary change in shaping gene expression variation in natural populations of extremophile fish. *Molecular Ecology*, *26*(22), 6384–6399. doi:10.1111/mec.14360

Pavey, S. A., Collin, H., Nosil, P., & Rogers, S. M. (2010). The role of gene expression in ecological speciation. *Annals of the New York Academy of Sciences*, *1206*, 110–129. doi:10.1111/j.1749-6632.2010.05765.x

Pfennig, D. W., Wund, M. A., Snell-Rood, E. C., Cruickshank, T., Schlichting, C. D., & Moczek, A. P. (2010). Phenotypic plasticity's impacts on diversification and speciation. *Trends in Ecology and Evolution*, *25*(8), 459–467. doi:10.1016/j.tree.2010.05.006

Pichancourt, J. B., & van Klinken, R. D. (2012). Phenotypic plasticity influences the size, shape and dynamics of the geographic distribution of an invasive plant. *PLoS ONE*, *7*(2), e32323. doi:10.1371/journal.pone.0032323

Pigliucci, M., Murren, C. J., & Schlichting, C. D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology*, *209*(12), 2362–2367. doi:10.1242/jeb.02070

Pigliucci, Massimo. (2005). Evolution of phenotypic plasticity: Where are we going now? *Trends in Ecology and Evolution*, *20*(9), 481–486. doi:10.1016/j.tree.2005.06.001

Price, T. D., Qvarnström, A., & Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1523), 1433–1440. doi:10.1098/rspb.2003.2372

R Core Team. (2017). R Core Team (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria. URL Http://Www.R-Project.Org/.*

Ragland, G. J., Almskaar, K., Vertacnik, K. L., Gough, H. M., Feder, J. L., Hahn, D. A., & Schwarz, D. (2015). Differences in performance and transcriptome-wide gene expression associated with Rhagoletis (Diptera: Tephritidae) larvae feeding in alternate host fruit environments. *Molecular Ecology*, *24*(11), 2759–2776. doi:10.1111/mec.13191

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers

differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. doi:10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. doi:10.1093/bioinformatics/btp616

Romero, I. G., Ruvinsky, I., & Gilad, Y. (2012). Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics*, *13*(7), 505–516. doi:10.1038/nrg3229

Schaum, E., Rost, B., Millar, A. J., & Collins, S. (2013). Variation in plastic responses of a globally distributed picoplankton species to ocean acidification. *Nature Climate Change*, *3*(3), 298–302. doi:10.1038/nclimate1774

Scheiner, S. M. (1993). Genetics and Evolution of Phenotypic Plasticity. *Annual Review of Ecology and Systematics*, *24*(1), 35–68. doi:10.1146/annurev.es.24.110193.000343

Schmid, M., Dallo, R., & Guillaume, F. (2019). Species' range dynamics affect the evolution of spatial variation in plasticity under environmental change. *American Naturalist*, *193*(6), 798–813. doi:10.1086/703171

Schwenke, R. A., Lazzaro, B. P., & Wolfner, M. F. (2016). Reproduction–Immunity Trade-Offs in Insects. *Annual Review of Entomology*, *61*(1), 239–256. doi:10.1146/annurev-ento-010715-023924

Shaw, J. R., Hampton, T. H., King, B. L., Whitehead, A., Galvez, F., Gross, R. H., … Stanton, B. A. (2014). Natural Selection Canalizes Expression Variation of Environmentally Induced Plasticity-Enabling Genes. *Molecular Biology and Evolution*, *31*(11), 3002–3015. doi:10.1093/molbev/msu241

Sikkink, K. L., Reynolds, R. M., Ituarte, C. M., Cresko, W. A., & Phillips, P. C. (2014). Rapid Evolution of Phenotypic Plasticity and Shifting Thresholds of Genetic Assimilation in the Nematode Caenorhabditis remanei. *G3: Genes|Genomes|Genetics*, *4*(6), 1103–1112. doi:10.1534/g3.114.010553

Silbermann, R., & Tatar, M. (2000). Reproductive costs of heat shock protein in transgenic *Drosophila melanogaster*. *Evolution*, *54*(6), 2038–2045. doi:10.1111/j.0014-3820.2000.tb01247.x

Smyth, G. K., Michaud, J., & Scott, H. S. (2005). Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics*, *21*(9), 2067–2075.

doi:10.1093/bioinformatics/bti270

Sokoloff, A. (1972). *The biology of Tribolium with special emphasis on genetic aspects I*. *Clarendon Press and Oxford Univ. Press, Oxford*. doi:10.1086/407731

Sokolova, I. M. (2013). Energy-limited tolerance to stress as a conceptual framework to integrate the effects of multiple stressors. *Integrative and Comparative Biology*, *53*(4), 597–608. doi:10.1093/icb/ict028

Sørensen, J G, Dahlgaard, J., & Loeschcke, V. (2001). Genetic variation in thermal tolerance among natural populations of *Drosophila buzzatii*: down regulation of Hsp70 expression and variation in heat stress resistance traits. *Functional Ecology*, *15*(3), 289–296. doi:10.1046/j.1365-2435.2001.00525.x

Sørensen, Jesper Givskov, Kristensen, T. N., & Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. *Ecology Letters*, *6*(11), 1025–1037. doi:10.1046/j.1461-0248.2003.00528.x

Stearns, S. C., & Kawecki, T. J. (1994). Fitness Sensitivity and the Canalization of Life-History Traits. *Evolution*, *48*(5), 1438–1450. doi:10.2307/2410238

Stoks, R., Govaert, L., Pauwels, K., Jansen, B., & De Meester, L. (2016). Resurrecting complexity: The interplay of plasticity and rapid evolution in the multiple trait response to strong changes in predation pressure in the water flea Daphnia magna. *Ecology Letters*, *19*(2), 180–190. doi:10.1111/ele.12551

Townsend, J. P., Cavalieri, D., & Hartl, D. L. (2003). Population genetic variation in genome-wide gene expression. *Molecular Biology and Evolution*, *20*(6), 955–963. doi:10.1093/molbev/msg106

Via, S. (1993). Adaptive Phenotypic Plasticity: Target or By-Product of Selection in a Variable Environment? *The American Naturalist*, *142*(2), 352–365. doi:10.1086/285542

Via, S., & Lande, R. (1985). Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity. *Evolution*, *39*(3), 505. doi:10.2307/2408649

von Heckel, K., Stephan, W., & Hutter, S. (2016). Canalization of gene expression is a major signature of regulatory cold adaptation in temperate Drosophila melanogaster. *BMC Genomics*, *17*(1), 1–14. doi:10.1186/s12864-016-2866-0

Whitehead, A., & Crawford, D. L. (2006). Neutral and adaptive variation in gene expression. *Proceedings of*

*the National Academy of Sciences of the United States of America*\r, *103*(14), 5425–5430. doi:10.1073/pnas.0507648103

Wray, G. A. (2007). The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics*, *8*(3), 206–216. doi:10.1038/nrg2063

Wund, M. A. (2012). Assessing the impacts of phenotypic plasticity on evolution. *Integrative and Comparative Biology*, *52*(1), 5–15. doi:10.1093/icb/ics050

Yampolsky, L. Y., Glazko, G. V., & Fry, J. D. (2012). Evolution of gene expression and expression plasticity in long-term experimental populations of *Drosophila melanogaster* maintained under constant and variable ethanol stress. *Molecular Ecology*, *21*(17), 4287–4299. doi:10.1111/j.1365-294X.2012.05697.x

Yampolsky, L. Y., Zeng, E., Lopez, J., Williams, P. J., Dick, K. B., Colbourne, J. K., & Pfrender, M. E. (2014). Functional genomics of acclimation and adaptation in response to thermal stress in *Daphnia*. *BMC Genomics*, *15*(1), 1–12. doi:10.1186/1471-2164-15-859

Yeh, P. J., & Price, T. D. (2004). Adaptive Phenotypic Plasticity and the Successful Colonization of a Novel Environment. *The American Naturalist*, *164*(4), 531–542. doi:10.1086/423825

Zhao, L., & Jones, W. A. (2012). Expression of heat shock protein genes in insect stress responses. *Invertebrate Survival Journal*, *9*(1), 93–101.

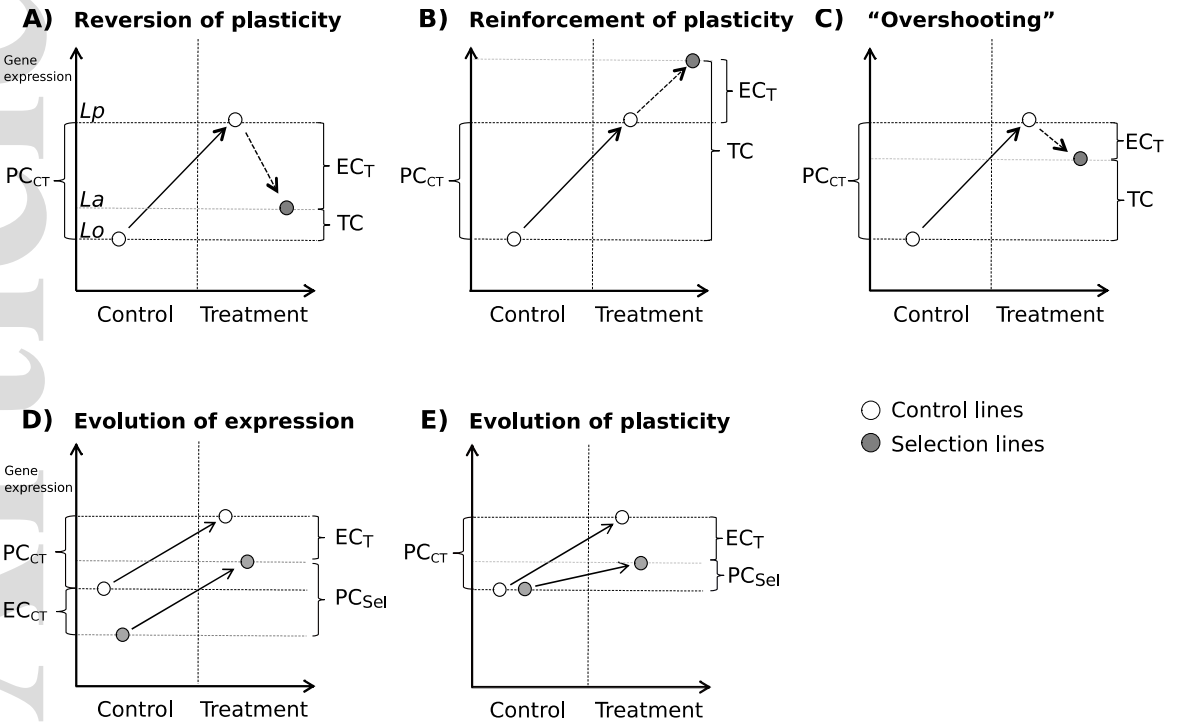



**Figure 1:** Possible relations between plastic change PC and evolutionary change EC. Gene expression levels of control lines (white) in the ancestral control condition represent the original stage Lo, expression in the treatment the plastic stage Lp. Expression levels of selection lines (grey) in the treatment give the adapted stage La. Arrows indicate the direction of PC (solid line) and EC (dashed line). PC can be in opposite direction to EC (reversion **(A))** or it can be in the same direction (reinforcement **(B))**. However, even if PC and EC are opposite to each other, PC can bring expression levels closer to levels of the adapted lines. In this case, the total change TC, (difference between expression levels of control lines in control conditions and selection lines in treatment) is larger than EC **(C)**. During adaptation, lines could have reached the optimum by either changing mean expression, i.e. shift in the intercept of the reaction norm **(D)** or by changing their plasticity, i.e. the slope of the reaction norm **(E)**. In case of a change in the mean, plastic changes of Control lines $PC_{CT}$ and selection lines $PC_{Sel}$ as well as observed evolutionary change $EC_T$ in treatment and in Control $EC_{CT}$ would be highly correlated because reaction norms (arrows) remain parallel **(D)**. If observed $EC_T$ in treatment is due to a change in plasticity only, no correlation between $EC_T$ and $EC_{CT}$ should exist, and the correlation between $PC_{Sel}$ and $PC_{CT}$ should be small **(E)**.

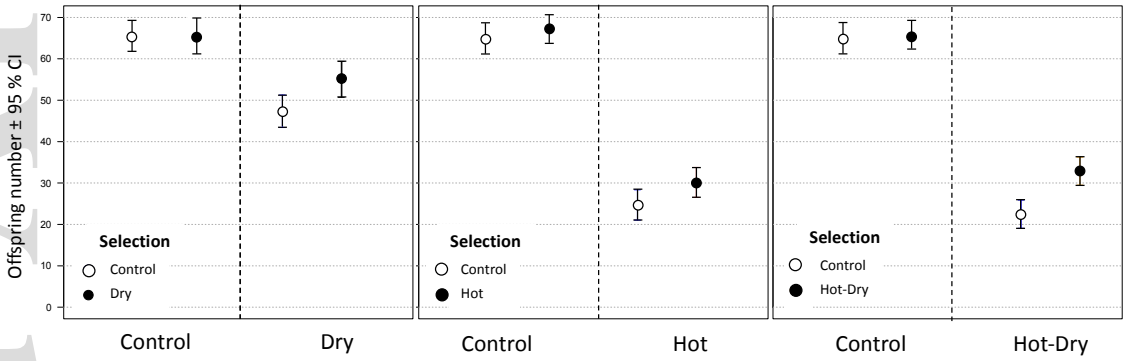



**Figure 2:** Offspring number of Control and selection lines under different conditions. Selection lines could adapt for 20 generations.

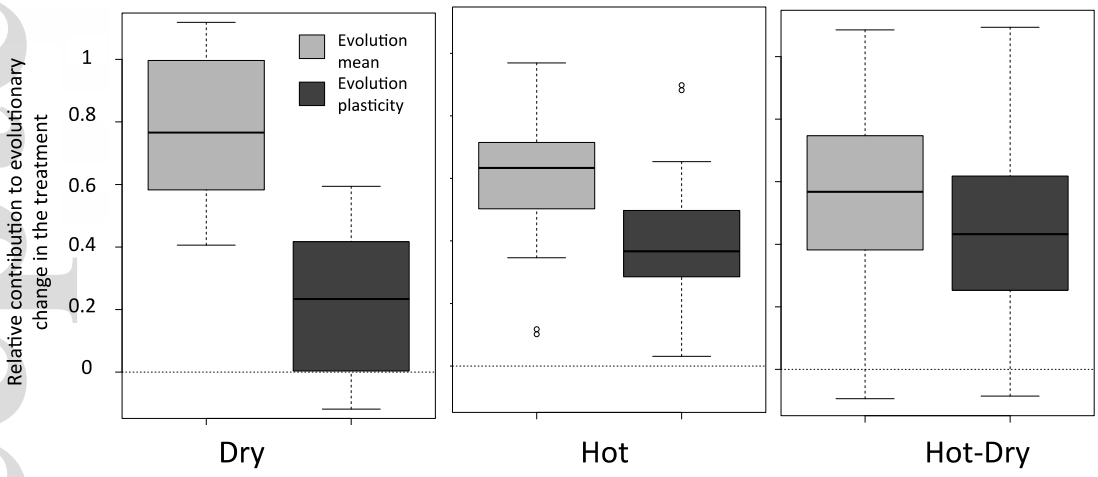



**Figure 3:** Relative contribution of changes in the mean (shift in the intercept of reaction norm) and changes in plasticity (different slopes of reaction norms) to evolutionary differences between control-lines and selection lines in the treatment. Only genes with significant differences in the DE analysis were included. Number of genes: Dry: 18, Hot: 25, Hot-Dry: 55.

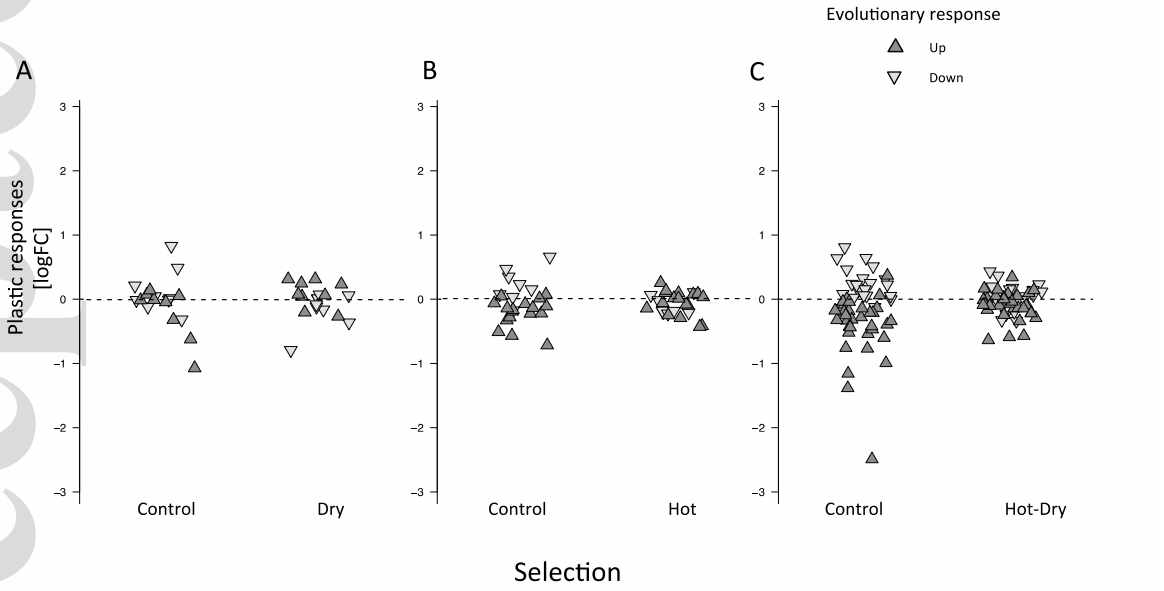
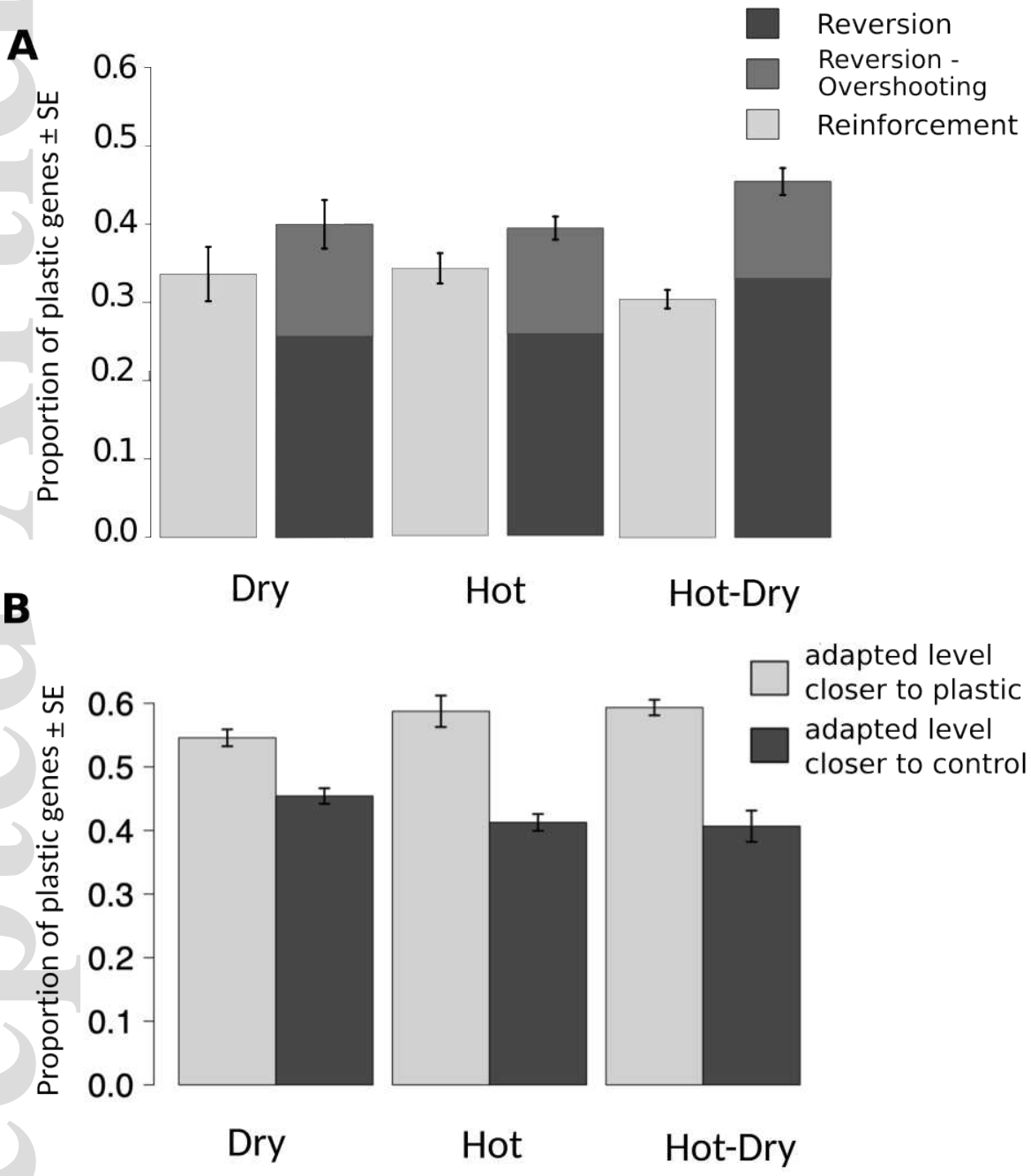



**Figure 4:** Plastic responses in control-lines and adapted selection-lines of genes showing significant evolutionary changes in expression in the treatments Dry **(A)**, Hot **(B)** and Hot-Dry **(C)**. Plastic responses in adapted selection lines are weaker (smaller logFC). Furthermore, plastic changes are mainly opposite to evolved changes. Genes that evolved to lower expression in the

968 treatments (represented in lightgrey) show positive plastic changes in control-lines, i.e. they are up-regulated. Genes that evolved
969 to higher expression levels (in darkgrey) are down-regulated in the non-adapted Control-Lines.

970




971

972 **Figure 5:** Comparison of plastic and evolutionary changes in gene expression in response to three environmental conditions. **A:**
973 Proportion of genes (average over selection lines) showing a reversion or reinforcement of ancestral plasticity during evolution.
974 Only genes exhibiting substantial plastic changes (difference between expression of Control-lines in Control (*Lo*) and treatment
975 (*Lp*) > 20 % of control expression) as well as evolutionary changes (absolute difference between Control- and selection-lines in
976 treatment > 20% *Lo*) are included. Even if plastic changes get reversed by evolution, they can bring *Lp* closer to *La* (here referred
977 to as 'overshooting', see also Figure 1C). In contrast, reversions where plastic responses move *Lp* further away from *La* (and *La* is

closer to the ancestral stage) indicate that they are maladaptive on the long term. **B:** Proportion of genes with expression levels after evolution (adapted stage *La*) closer to original levels (*Lo*, i.e. Control-lines in control conditions), or closer to plastic levels (*Lp*, i.e. expression levels of Control-lines in treatment). Only genes exhibiting substantial plastic changes (|*Lp-Lo*| > 20% of *Lo*) were used for this analysis. Results shown here include four selection lines in Dry, five in Hot and seven in Hot-Dry. Results for each selection line separately as well as results for using different thresholds for defining substantial plastic and evolutionary responses can be found in Supporting Information Figure S5, Table S4.

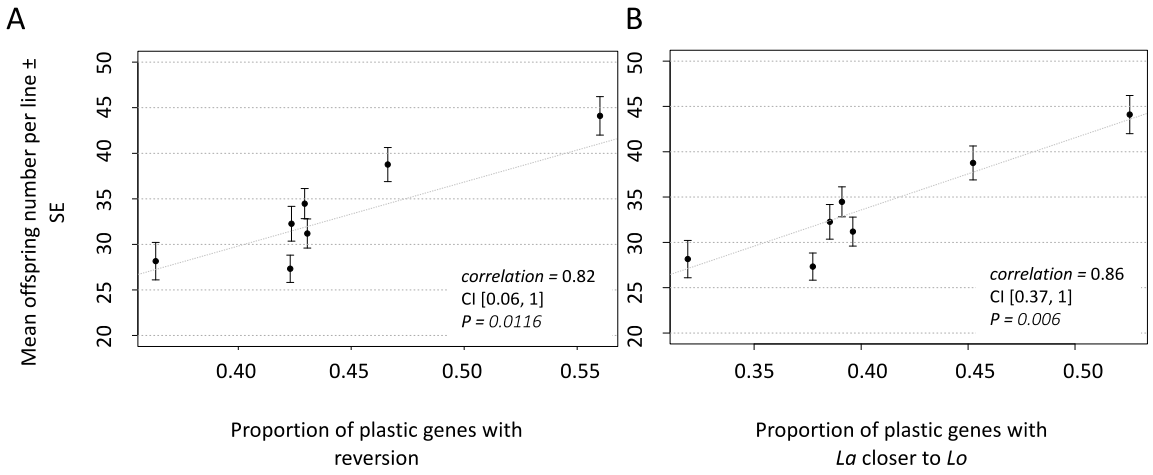



**Figure 6**: Relationship between plastic and evolved changes in gene expression in response to hot-dry conditions. Expression levels of Control-lines in control conditions represent the ancestral stage Lo, Control-Lines in hot-dry conditions the plastic stage Lp and adapted Hot-Dry-lines in hot-dry conditions the adapted stage La. **A:** Relationship between the proportion of reversed plastic responses and mean fitness (=offspring number) per Hot-Dry line. **B:** Relationship between proportion of genes with La closer to Lo and mean fitness. Only genes exhibiting substantial plastic changes (|Lp-Lo| > 20% of Lo) as well as evolutionary changes (La-Lp| > 20% of Lo ) were used for analysis. P-values of the spearman correlations were obtained by 10,000 permutations. 95 % Confidence intervals are based on a non-parametric bootstrap test. Results for Dry and Hot (with four and five selection lines only) are shown in Figure S8.

**Table 1:** Number of sequenced replicate lines and individuals per selection and treatment, which were used for this study. Selection lines could adapt to conditions for 20 generations. Control (CT) conditions: 33°C, 70% relative humidity (r.h.); treatments: Dry (D): 33°C, 30% r.h.; Hot (H): 37°C, 70% r.h.; Hot-Dry (HD): 37°C, 30% r.h.

| Selection | | Conditions | | | |
|---|---|---|---|---|---|
| | Total | CT | D | H | HD |
| CT | | | | | |
| **Number lines** | **7** | 7 | 5 | 5 | 7 |
| **Number Ind.** | **89** | 25 | 20 | 19 | 26 |
| D | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | **Number lines** | **4** | 4 | 4 | – | – |
| | **Number Ind.** | **32** | 16 | 16 | – | – |
| H | | | | | | |
| | **Number lines** | **5** | 5 | – | 5 | – |
| | **Number Ind.** | **40** | 20 | – | 20 | – |
| HD | | | | | | |
| | **Number lines** | **7** | 7 | – | – | 7 |
| | **Number Ind.** | **56** | 28 | – | – | 28 |

**Table 2:** Significantly differently expressed genes. For evolved differences, control-lines and selection-lines are compared within condition (e.g. Dry-lines vs Control-lines in Dry). Plastic response gives the number of genes that changed expression between control and treatment conditions (e.g. Dry-lines in Dry vs Dry-lines in control). Different plasticity gives the number of genes with significant different plastic responses in lines from different selection regimes. Acronyms in brackets refer to Figure 1. Analysis was conducted using the R package limma (Ritchie et al., 2015).

| | | Dry | Hot | Hot-Dry |
|---|---|---|---|---|
| **Evolved difference in Control ($E_{CT}$)** | down | 3 | 0 | 6 |
| | up | 1 | 9 | 4 |
| | **total** | **4** | **9** | **10** |
| **Evolved difference in treatment ($E_T$)** | down | 9 | 8 | 22 |
| | up | 9 | 17 | 33 |
| | **total** | **18** | **25** | **55** |
| **Plastic response Control-Lines ($PC_{CT}$)** | down | 209 | 1765 | 2451 |
| | up | 156 | 1714 | 2200 |
| | **total** | **365** | **3479** | **4651** |
| **Plastic response Selection-Lines ($PC_{Sel}$)** | down | 28 | 1417 | 1649 |
| | up | 21 | 1381 | 1470 |
| | **total** | **49** | **2798** | **3119** |
| **Total change (Selection-lines in treatment vs Control-lines in Control) (TC)** | down | 154 | 564 | 1016 |
| | up | 129 | 628 | 1029 |
| | **total** | **283** | **1192** | **2045** |
| **Different plasticity** | **total** | **0** | **1** | **4** |

**Table 3:** P-values obtained from permutation tests (10,000 permutations). Samples were randomly assigned to either control or treatment selection and differential expression analyses were repeated. Significance was assessed by calculating the proportion of permutations with more extreme values than the observed one. Control (CT) conditions: 33°C, 70% relative humidity r.h. Conditions in treatments: Dry (D): 33°C, 30% r.h.; Hot (H): 37°C, 70% r.h.; Hot-Dry (HD): 37°C, 30% r.h.

| | **Selection** | | |
|---|---|---|---|
| | **D** | **H** | **HD** |
| Control-lines have more genes with significant plastic responses compared to adapted selection lines | 0.2981 | 0.3037 | **0.0325** |
| Magnitude of plastic response is higher in Control-lines | **0.0493** | 0.1676 | **0.0491** |
| Number of genes with significant differences in expression levels is higher in treatment than in CT conditions | **0.0095** | **0.0003** | **0.0031** |
| Differences in expression levels in CT conditions and treatment are correlated | **< 0.0001** | **<0.0001** | **< 0.0001** |